

ISSUE 11 JANUARY 2010

Big Picture

BRINGING CUTTING-EDGE SCIENCE INTO THE CLASSROOM

GENES, GENOMES
AND HEALTH

The wonder of you

Exploring your genetic identity

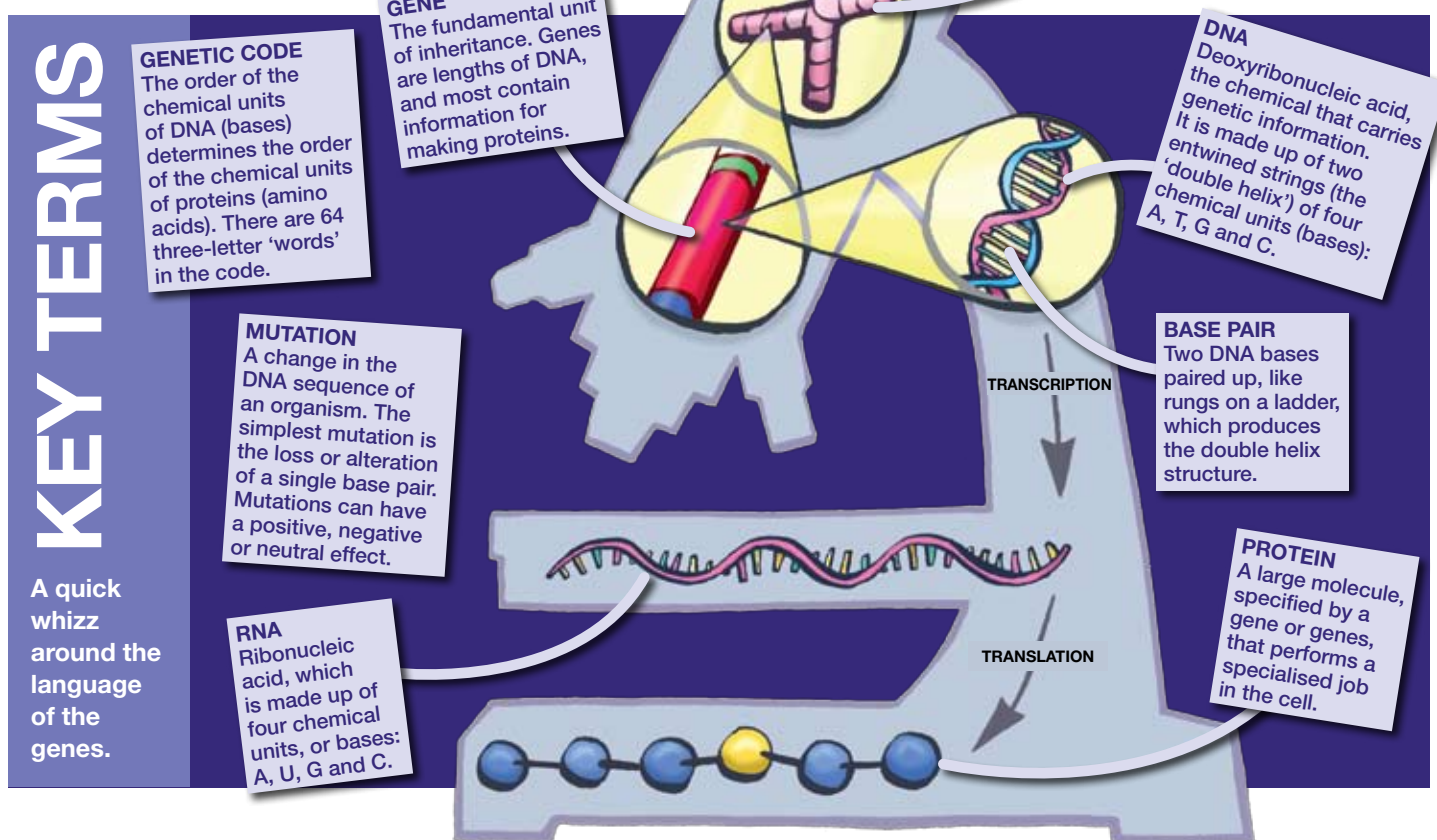
A resource for teachers and learners
FREE

- Inside the human genome
- The genetic basis of health and disease
- Regulating the use of genetic information
- What DNA studies tell us about the past

Big Picture on genes, genomes and health

Beyond the human genome

Our genes play a key part in making us who we are, but how can science help us unpick and understand our genetic identity? Since the working draft of the human genome was unveiled in 2000, mind-boggling progress has been made in our ability both to sequence a genome accurately and quickly, and to process and understand the huge amount of data these activities produce. What do these developments mean for each of us – our health, identity, and the world we live in – now and in the future?



All in sequence

Why is determining an organism's genome sequence important?

Most types of cell in an organism contain a complete copy of its genome. The organisation is quite complicated, but the simplest fact about any genome is that it is a collection of DNA sequences – long strings of the chemical 'letters' A, T, G and C (adenine, thymine, guanine and cytosine) in a particular order.

Learn to read an organism's genome sequence, and compare it with that of other organisms, and it can tell you lots of different things.

The human genome sequence contains a wealth of information about human biology, in both health and disease. Our DNA is a

window on to evolution and recent human history – including the migration of people around the world. DNA can also provide clues in police work, from crime scene samples, and can help to identify disaster victims.

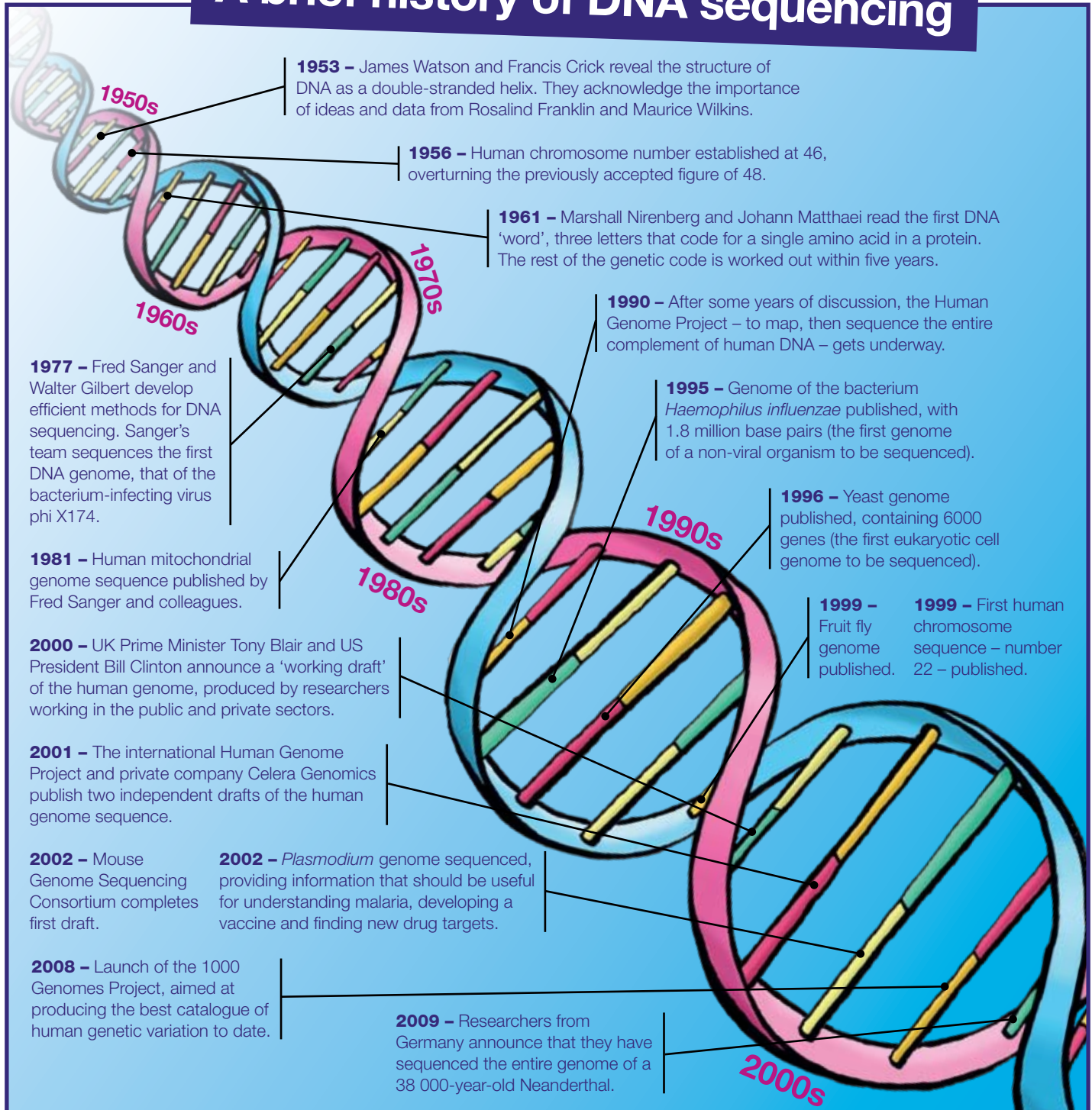
The genome sequences of other species have many other uses. The genomes of organisms used in farming, from rice and wheat to pigs and cattle, are being sequenced to help to breed improved strains. But the vast majority of the many hundreds of genomes already completed are from bacteria. Some are species that cause diseases in people, as well as in agriculturally important animals or plants.

Others are important for maintaining health, or have potential use in the industrial production of biologically active chemicals and enzymes.

Genomic information is used to track harmful bacteria such as those that cause infection in hospitals, as well as to aid the development of new drugs. New flu strains have their genomes read quickly to understand how the virus spreads and to speed up vaccine production.

Knowledge of genome sequences also speeds up developments in biotechnology, and is finding uses in tracking biodiversity and policing trade in protected species.

A brief history of DNA sequencing



\$1000 genome

Could getting our genomes sequenced be an affordable option in the future?

Sequencing is still getting cheaper. In August 2009, US engineer Stephen Quake announced his company's 'Single Molecule Sequencer', a million-dollar machine that allows three people

to sequence a human genome in four weeks. The bill? Around \$50 000. Illumina, a company that makes sequencing machines that are used in many of the large genome-sequencing centres, charges the same for a new personal genome service.

A \$1000 genome analysis would open the way to routine use in medicine, although it is still probably some years away. To help speed things up, a US-based foundation is offering a \$10 million award. The X Prize for Genomics will go to whoever builds a machine that can sequence 100 human genomes in ten days, with no more

than one error in every 100 000 bases, for no more than \$10 000 per genome.



ON THE WEB

CH-CH-CH-CHANGES

Discover how gene sequencing technologies have developed since the 1980s.

www.wellcome.ac.uk/bigpicture/genes



Introducing the genome

Many of the scientists working to decode the human genome were surprised to find far fewer genes than expected. Subsequent work continues to uncover interesting features of the human genome, which has proved to be so much more than just a collection of genes. What makes up the human genome? How does it compare with those of other organisms?

Great expectations

The draft human genome held many surprises for researchers.

Printing all 3 billion letters of the human DNA sequence would fill 200 telephone directories. When the 'book of life' was fully decoded, it turned out to be full of surprises.

Most scientists had assumed that humans had around 100 000 genes. Wrong: the true number is nearer 25 000. So does this mean that we're much simpler genetically than rice, for example, which has around 50 000 genes? Not necessarily – the human genome harbours some organisational secrets.

Human cells make perhaps three times as many proteins as they have genes. This

is achieved through 'alternative splicing', which basically means taking genes, or their messenger RNAs, in bits and reshuffling them. The result is rearranged proteins, with different structures.

This makes it harder to define what a gene is. It is no longer a piece of DNA that always does the same thing – 'one gene, one enzyme', as biologists used to say. Many genes are shape-shifters, making one product then another. What's more, many genes can 'multitask' – playing different roles in different tissues, or at different life stages.

The versatility of genomes is still being revealed. Recent findings indicate that some 'jumping genes', which can move around the genome, are especially active during brain development, for example – which might allow the brain to generate new types of neuron. Nor is the human genome sequence quite complete. Certain regions of chromosomes, usually with highly repetitive sequences, are hard to nail with current technologies. There are also still a few gaps in the coverage of actual genes, though these amount to under one per cent of the total.



James Watson, co-discoverer of DNA's double helical structure.

Adrian Brooks

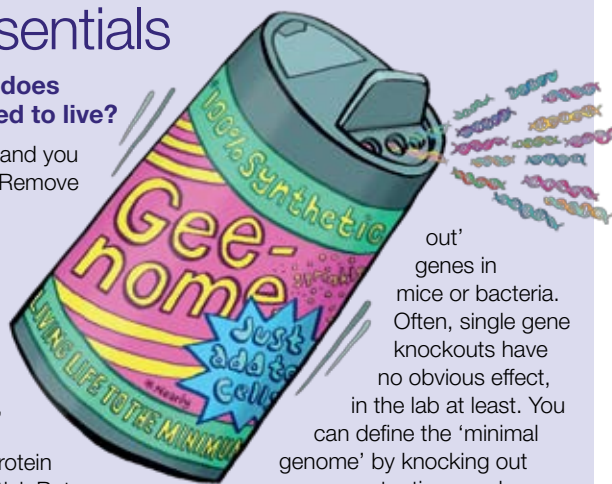
Bare essentials

How few genes does an organism need to live?

Turn off the sat nav and you can still drive a car. Remove the fuel pump and you won't be going anywhere. Similarly, even simple organisms carry genes they can do without, at a pinch. Some genes, such as those that code for the cell's protein factories, are essential. But how few genes can an organism get away with?

Viruses don't count here because they need – for example – a sneezing host's cells to reproduce. But some independently living bacteria are pretty light on genes. One of the simplest found so far is *Mycoplasma genitalium*, which – as its name suggests – lives on the genitals of primates. Its tiny genome, sequenced in 1995, has around 500 genes.

Genetic engineers can now 'knock



out' genes in mice or bacteria. Often, single gene knockouts have no obvious effect, in the lab at least. You can define the 'minimal genome' by knocking out genes one at a time, and checking whether the organism survives. The bacterium *Bacillus subtilis* is able, if supplied with good nutrients, to manage without all but 271 of its 4100 genes.

More ambitious is to make your own genome from scratch, then stick it into an empty cell and see whether it works. Craig Venter is trying this approach using *Mycoplasma*. If he succeeds, this 'synthetic organism' might then be developed for industrial applications such as creating biofuels.

More than genes

The genome is far from just a collection of genes.

Genes take up only a small part of the genome: protein-coding sequences account for only around 1.5 per cent of human DNA.

Biologists used to regard the other regions (which don't code for proteins) as 'junk'. Now though, these regions are slowly yielding their secrets. Some of the non-coding DNA regions are probably parasitic bits and pieces that are along for the evolutionary ride. But at least 10 per cent and perhaps a lot more of the whole genome is transcribed into RNA at some time. The role these 'extra' RNAs play is still being teased out. They are often short-lived and some are much smaller than mRNAs.

RNA, because it has a single strand of bases, can bind to a

complementary RNA or DNA. This seems to make it ideal for lots of the detailed regulation of gene action. Small RNAs that bind to mRNA or DNA can stop a protein being made or tag the mRNA for disposal, for example. The newfound versatility of RNA is giving it a starring role in the workings of the cell – but every RNA is still produced from reading a stretch of code in the genome.

Just how many RNAs there are will become clearer when the US-led Encyclopedia of DNA Elements (ENCODE) project is complete. Researchers are looking at more than 40 different kinds of human cell, to make sure they catch all the parts of the human genome that get read out into RNA.

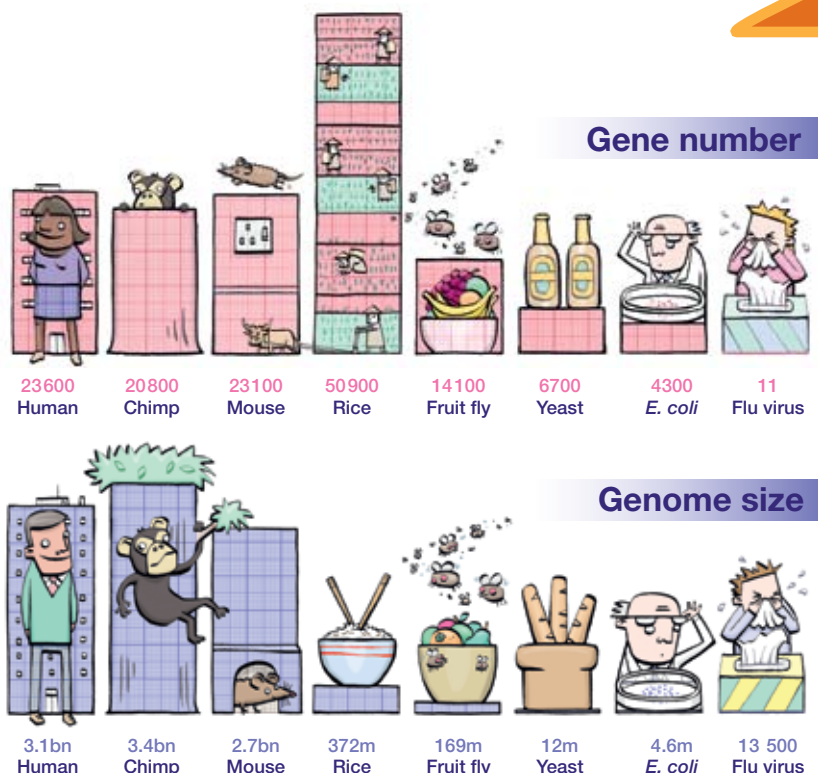
Measuring up

How does the human genome compare with those of other organisms?

FAST FACT

It would take 9.5 years non-stop to read aloud a person's genome base by base.

Source: genomics.energy.gov



Gene number refers to protein-coding genes. Genome size measured in base pairs.
 Rice: Japonica subspecies; yeast: *Saccharomyces cerevisiae*; *E. coli*: K-12.
 All figures from Ensembl (www.ensembl.org), except rice (rice.plantbiology.msu.edu),
E. coli (*Science* 1997;277:1453–62) and flu virus (*Nature* 2005;437:1162–6).

Now we have whole genomes from a range of organisms, comparing them is a powerful way to investigate what makes each creature distinctive. The simplest measure is to count genes and check which are there. Do that and you find, for example, that the tiger puffer fish (*Fugu rubripes*), the first backboneed creature to have its genome sequenced after humans, has many genes in common with us. On the other hand, around a quarter of human genes have no equivalents in the fish, and we have almost ten times as much DNA.

Animals without backbones, such as fruit flies and nematode worms, share very many genes with humans too.

Researchers are still poring over the chimpanzee genome, first published in draft in 2005. This confirms that the chimp is our closest living relative, with a genome 98.8 per cent identical to ours.

But what does this figure mean? For very closely related organisms, such as humans and chimps, we can align most of one organism's genome with that of

the other. This reveals simple differences and allows us to calculate a percentage similarity. Such a comparison becomes less meaningful with organisms that are less closely related, e.g. humans and nematode worms, because less of the two genome sequences can be aligned together.

So, even for our closest relative, the crude measure of 98.8 per cent similarity conceals important differences. Apart from the precise sequence of genes there are some differences in the overall set of genes. Also, the time when particular genes switch on and off during development does not show up, and this is very different for some genes – especially ones that affect the brain. These subtle differences in gene regulation and expression are likely to underlie the most important differences between closely related species.



ON THE SLIDE

Try our interactive slider to explore how similar, or not, we are to a number of organisms.

ON THE WEB

www.wellcome.ac.uk/bigpicture/genes

Whose genes?

The reference human genome combines the DNA of several people.

The original of 'the' human genome was a reference sequence, compiled by analysing DNA from blood samples donated by a handful of anonymous volunteers. It stands as a comparison for genomes analysed later on.

That sequence covers the nuclear DNA – that contained in the cell's nucleus – over 99 per cent of the total. The mitochondrion, the tiny organelle that acts as the cell's energy generator, has its own separate genome, with 16 500 base pairs and just 37 genes. Every cell has hundreds of copies of this DNA, and many different samples have been sequenced. For more on mitochondria, see page 8.

Since the Human Genome Project was completed, a few named individuals have had their genomes analysed. The first two were Craig Venter (the driving force behind the private genome project) in 2007, and DNA pioneer James Watson the following year. Genome analyses have since been published for individuals from three distinct geographic regions: Korea, China and Nigeria (in the latter case, one of the Yoruba people).

Meanwhile, the 1000 Genomes Project, launched in 2008, will refine the reference sequence with information from a larger number of anonymous samples. The project, involving researchers in the UK, the USA and China, is cataloguing variations between individuals that occur in at least 1 per cent of the population to pinpoint differences in DNA relevant to health and disease.

Clockwise from left: A Han Chinese man, a Korean woman, a Yoruba man from Nigeria. The genome analyses published so far include those for individuals from China, Korea and Nigeria.



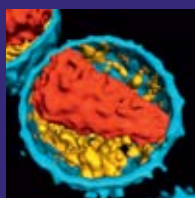
In sickness and in health

Genetics plays a part in many diseases. In a few conditions, such as cystic fibrosis, a particular mutation in a single gene is enough to cause disease. In the majority of diseases, though – including common disorders such as diabetes and asthma – the influence of our genes can be much more subtle. What do we know now about the genetic basis of disease? How has the progress in genome sequencing informed the way we categorise, diagnose and treat conditions?

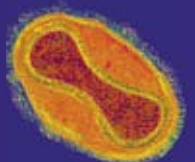
Can you resist?

Genetic variations play a part in disease resistance.

Some gene variants are good news. If you are of northern European descent, for instance, there is a 1 per cent chance that you carry two mutated copies of the *CCR5* gene, which means that key cells in your immune system lack working copies of the *CCR5* receptor. This receptor is what HIV latches on to when it attacks those cells, so a lack of working receptors provides defence against HIV. About 10 per cent of the same population lack one working copy, and have partial protection. Why did this happen? One possibility is that the gene mutation arose when smallpox was an everyday hazard, and persisted because it helps to prevent that scourge as well.



Stephen Fuller

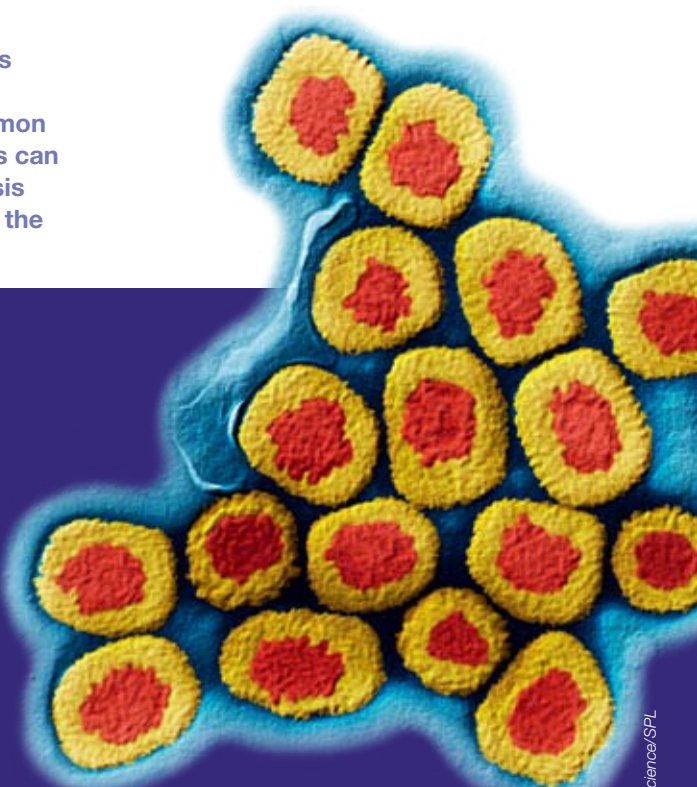


AMI Images/SPL

Top left: Internal structure of an HIV particle, including the RNA-containing core (red).

Bottom left: Smallpox virus particle, including its DNA genome (red).

Right: Smallpox viruses. Smallpox was eradicated in the 1970s by a global vaccination programme, but is the virus's once-widespread presence the reason some people are protected from HIV infection today?



Eye of Science/SPL

Nature and nurture

Most conditions involve an interaction between genes and environment.

Almost everything about health is affected by genes. Most often, the effect of an individual genetic variation is small, and is usually influenced by additional external factors: diet or exercise, exposure to a virus, bacterium or radiation, or a more general challenge such as heat stress or exhaustion.

For example, people who develop the cancer mesothelioma have almost always been exposed to asbestos, the fibres of which lodge in the lungs. Some people exposed will be at a higher risk because of minor genetic differences. Which genes matter is still being investigated.

Some genetic differences have mixed costs and benefits.



People in protective workwear removing asbestos.

There are a number of known mutations in the gene for the blood's oxygen-carrying protein haemoglobin. Carrying one altered copy of the gene can help to protect against infection by the parasite that causes malaria. However, carrying two altered copies (and hence no normal haemoglobin) can lead to diseases including sickle-cell anaemia or thalassaemia.

Cancer cells have aberrant genomes, usually as a result of a series of genetic changes that happen as the disease develops. In the first step, normal cells may be exposed to chemicals that damage DNA or interfere with its repair.

For more on the influence of genes and environment on disease, see page 8.

Seeing into the future

Why finding a 'disease gene' is just the beginning of finding a treatment or cure.

Once a gene is linked firmly to a particular condition, it may lead to a genetic test, perhaps before birth, to assess disease risks. Going beyond that can prove a lengthy effort.

Cystic fibrosis is now one of the better-understood genetic diseases. We know the gene, all the ways it can be altered, the protein that is affected and what job the protein does in cells.

There are tests available for the condition, helping carriers of the mutation to make informed choices about having children, for example (see page 13). Yet 20 years after the gene was found, treatments that build on this remain elusive.

Therapy has improved over those decades, and people with the condition now typically live a great deal longer. But no treatments have yet come through that are targeted at the protein – the cystic fibrosis transmembrane conductance regulator (CFTR) – whose defects cause lung congestion, digestive problems and damage to the pancreas.

However, there are promising results in trials of drugs that may restore the function of some altered forms of CFTR. There are also early trials underway with gene therapy, building on earlier research in this area, to give people a dose of the normal gene so they can make the right form of the protein. The cystic fibrosis saga does not suggest that unravelling disease genes is not helpful, just that the benefits take time, and a lot more hard work from researchers.

Jennifer Grannell, who was diagnosed with cystic fibrosis when she was two, with some of the medications she uses.



Cystic Fibrosis Trust



What do we know now?

How is current research investigating the genetic basis of disease?

Access to data about the entire genome is increasing knowledge of many diseases, but also reinforcing awareness of how complex cells, tissues and bodies are. Although the press may still report a 'gene for' things like arthritis or depression, studies almost always show that there are multiple factors involved.

One powerful research approach is to use genome-wide association studies. These involve going through the entire genome looking for statistically significant links between genetic variation and a disorder. Usually, the larger the population studied, the more genetic links that appear. But each link contributes only a tiny fraction of susceptibility to the disease. The method does not tell you what each

gene region does, so each one has to be looked at more closely for clues to the underlying mechanisms of disease.

Such studies can produce strong findings. Several teams, including one led by Professor Andrew Hattersley at the Peninsula Medical School in Plymouth, have identified genes that affect the risk of developing type 2 (late-onset) diabetes. This disease is caused by the body not responding properly to the hormone insulin, and further investigation is helping to tease out the complex regulation of responses.

Some genome-wide association studies have proved less fruitful so far – for example, those looking for genes linked to some psychiatric illnesses, including schizophrenia and depression.

Getting personal

Is the future of pharmaceuticals really all me, me, me?

Gene tests could help doctors to identify the best treatment for each patient. This is important because some gene variants affect how the body breaks down different drugs, for example. Giving the right drug to the right patient at the right dose, when there is a choice, should increase effectiveness and reduce adverse reactions.

A more complete knowledge of how genetics affects our responses to drugs – pharmacogenetics – could eventually lead to an era of personalised medicine.

Examples so far include identifying variants in a key enzyme complex in the liver known as cytochrome P450, which metabolises drugs. Some types of P450 reduce the effectiveness of antidepressants, while others diminish people's responses to painkillers such as codeine. Another example is Herceptin, a drug given to women with breast tumours if they have an overactive gene for a receptor known as HER2. They benefit from Herceptin, an antibody that blocks the receptors on the cancer cells, whereas women without the receptor do not.

Personalised medicine has economic implications. Drug companies spend heavily on research to find compounds that will be prescribed for thousands of people. Pharmacogenetics might mean each drug is used by a smaller number of patients, although each of them would benefit more.

Women with breast cancer, their relatives and supporters on a march in Cardiff to demand the end of a Herceptin 'postcode lottery', where people in different areas have different access to particular treatments.



Johnny Green/PA Archive/Press Association Images

Single-gene and chromosome disorders

Some conditions are caused by genetic 'abnormalities'.

Some conditions are so tightly tied to DNA mutations that they can be called genetic diseases. In a few cases, inheriting just one copy of a mutated gene (i.e. a 'normal' version from one parent and a 'mutated' version from the other) can lead to disease. These are known as dominant conditions, and include Huntington's disease, an incurable condition that causes gradual deterioration of the brain (for more on this see page 15).

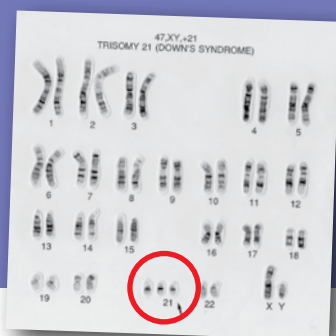
Recessive conditions, such as cystic fibrosis, are a

larger group of genetic disorders that occur only if you inherit two copies of a mutated gene (i.e. one from each parent).

There are also conditions due to larger changes in the chromosomes. Down's syndrome, for instance, is caused by an extra copy of chromosome 21.

These diseases can be more complicated than they may seem at first glance. The protein affected in cystic fibrosis can be altered by more than a thousand different mutations. Some have almost no effect, while others cause severe lung and stomach problems. The extent of symptoms can also be affected by other genes, which modify the effect of the cystic fibrosis gene. People

with Down's syndrome differ on a whole range of characteristics, which depend on factors including how much of chromosome 21 is present in three copies.



GOT THE BUG?

From protecting orange trees to fighting the plague, find out what we can learn from sequencing the genomes of all kinds of bacteria and viruses.

ON THE WEB

NO MORE INSULIN

Watch a video to find out how the discovery of a genetic mutation that caused neonatal diabetes has led to a new treatment that's transformed the lives of patients and their families.

www.wellcome.ac.uk/bigpicture/genes

Genes and the wider world

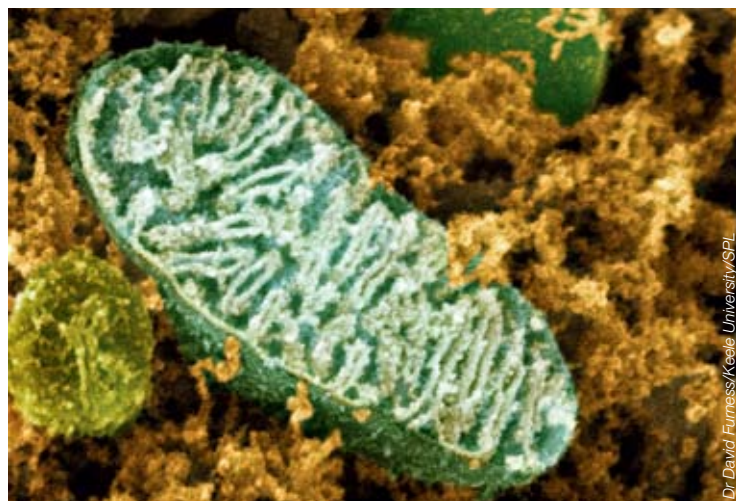
Just as researchers conduct studies to unpick the genetic basis of disease, they must also consider how our lifestyle and surroundings interact with our genome to shape our current and long-term health. It's not all about the DNA found in our cells either – studies of the billions of microorganisms that call our bodies home are yielding fascinating insights into everything from the bacterial causes of medical conditions to the history of human migration.

Kelly Cline/Stockphoto

FAST FACT

In an average meal you eat around 150 000 km of DNA.

Source: Iowa State University Office of Biotechnology



Dr David Furness/Keble University/SPL

Beyond nuclear

Mitochondria mean that there's a bit more of mum in all of us.

There's another bacterial genome, or some of one, in all of our cells – but it lies in the energy-generating organelles, the mitochondria (one shown above). These vital accessories are almost certainly descendants of symbiotic bacteria that colonised larger cells more than 2 billion years ago. Most of a mitochondrion's essential functions depend on genes in the cell's nucleus (which have been transferred there over millions of years), but mitochondria retain and use a separate genome.

Mistakes in copying the mitochondrial DNA seem more common than mistakes made when copying nuclear DNA. The

mutations this causes can lead to mitochondrial diseases, an often perplexing set of conditions that can have widely varying symptoms. Not all mitochondria may be affected, and the consequences can be different in different tissues.

When a sperm cell unites with an egg, it brings some mitochondria from your dad. These are eliminated soon after fertilisation, however, so the hundreds of mitochondria in each of your cells are descended from your mum. This unusual maternal pattern of inheritance for these few mitochondrial genes is important in studies of evolution.



TEACHERS' PACK

Check out free videos, animations and more.

ON THE WEB

www.wellcome.ac.uk/bigpicture/genes



Understanding the links

How do researchers tease apart genetic and environmental factors?

Before the genome era, twin studies helped researchers to assess the relative influences of genes and environment – without knowing details of particular genes. Researchers can record what happens to lots of pairs of identical twins, and similar numbers of non-identical, same-sex twins, when each pair is raised together in a similar environment. If some characteristic differs more in non-identical twins, this is likely to be because of some genetic effect.

Better, but more difficult to arrange, is to look at identical twins who have been raised in different households (and, therefore, different environments). If they nevertheless turn out very alike, the likeness will be, at least partly, due to genes shaping their development.

Identifying the genes involved requires a different kind of study. For example, Professor Nick Wareham at the UK Medical Research Council's Epidemiology Unit in Cambridge is leading a study to understand the complex effects of diet, exercise and other behaviours on the risk of developing type 2 diabetes, and how this risk is modulated by genes.

The researchers are using a mega-database of 350 000 people across Europe, which was assembled for a study of diet and health originally designed to investigate causes of cancer. The records are so detailed that they will allow the Cambridge team to pull out 10 000 people with diabetes and match them with 10 000 diabetes-free controls. This will provide a large enough group to get good indications of which genes and lifestyle differences are important contributors to diabetes risk. The results will feed into health policies to tackle what is becoming one of the most common conditions of middle age.

Indeed, there could be scope to use genetic information more widely to develop population-based interventions and campaigns, where preventative strategies can be targeted at the subgroups most at risk of particular conditions.



The genomes in us

Like it or not, our bodies are teeming with billions of microorganisms.

We carry many more genomes than just our own. Thousands of different kinds of bacterium live on and in humans, and there are perhaps ten times as many bacterial cells present as there are cells in our bodies. The Human Microbiome Project is an international effort uniting the labs sequencing the genomes of these passengers, both welcome and unwelcome. The total number of genes involved is probably more than a hundred times as big as in the actual human genome.

Meanwhile, there is a growing number of studies of what these bacteria do for us, or to us. Some bacteria are essential, such as those needed for digestion. Some are harmful. Some can be benign but turn nasty.

An example of the last kind is *Neisseria*

meningitidis, which lives harmlessly in the nose and throat of about four people in ten. But the bacterium, whose genome was sequenced in 2000, also causes around half a million cases of meningitis every year worldwide.

Much current research focuses on the bacteria in our intestines, most of which cannot live anywhere else. There is a whole ecosystem in there. Variations in this vast population – the gut microbiota – seem to affect the risk of obesity, as well as subtler responses to our diet. Other effects of microbial genomes include influence on allergies and asthma. A growing number of serious medical conditions formerly thought to be microbe-free, from ulcers to heart disease, turn out to have some bacterial culprits.

Future plans

Long-term population studies help researchers plan for future discoveries today.

The effects of the long-term interaction between genes and environment weave through a life story, and the medical consequences often appear late in life. That means waiting a long time for research results, too. Nowadays, the collection of DNA samples from people who enrol in studies is done alongside plans to keep the samples for many decades, and to make sure that any new tests that appear along the way can still be applied to them.

One far-sighted study in the west of England began in 1991. The Avon Longitudinal Study of Parents and Children (ALSPAC) signed up more than 14 000 mothers before their babies were born. The children have been followed ever since – and, as young adults, some are on the advisory panel that oversees the project. The research explores a range of social and medical issues. Findings so far include that eating oily fish in pregnancy improves the child's eyesight, and that children growing up in very hygienic homes are more likely to get asthma.

Larger still is UK Biobank, a national study that is recruiting 500 000 people aged 45–69. They will have a health screening interview, donate blood samples and agree to allow their future medical records to be shared with the researchers. The big numbers should help to reveal the numerous factors, each with small effects, that contribute to the development of a disease – although there are already ideas in discussion to combine the data with even larger studies from, for example, China.



Epigenetics

Is it through epigenetics that environment affects our DNA?

The gene–environment story is complex enough, but has acquired a new twist in recent years with growing evidence that environment can actually modify genes – after a fashion.

The modification is indirect. DNA sequences stay the same, but there are changes in the set of chemical tags that the genes, or the proteins bound to them, carry. These tags, most often simple methyl (CH_3) groups, affect gene expression – usually by preventing transcription.

Often, the gene affected is itself one that controls another

genetic switch or switches.

This means that there may be a cascade effect, where one small change leads to larger adjustments in gene expression.

Study of this kind of secondary modification of the genome is known as epigenetics, and the whole set of tags is the epigenome. The epigenome is a bit like the personal settings that gradually customise the operating system on a computer. Most, but not all of it, is 'reset' in sperm and egg. But some epigenetic changes can pass down to the next generation.

There is good evidence that a mother's diet during pregnancy and even just after birth can affect expression of some of a baby's genes – in mice and,

probably, in humans. And some of those changes will still be there in the grandchildren. So a famine, say, will leave traces down the generations in the way survivors' children metabolise their food.

A group of researchers based at the London School of Hygiene and Tropical Medicine is currently looking more closely at these effects in mothers and children in the Gambia. Those who live on the land there have more and fresher food at harvest time than during the rest of the year, so the researchers will monitor the effects on babies born in different seasons. The study will look at the exact links between specific nutrients in the mothers' blood samples and DNA methylation in their offspring.

Identity: your genes and you

Genetically speaking, humans are all remarkably similar. But most of us, including identical twins, see ourselves as highly individual people with our own personality, appearance and sense of identity. How far do our genes go in making each of us look and behave as an individual? What can studies of our genes tell us about the history of humanity and – more personally – how closely related we are to other people?

It's all history now

How DNA studies help to trace the origins of modern humans.

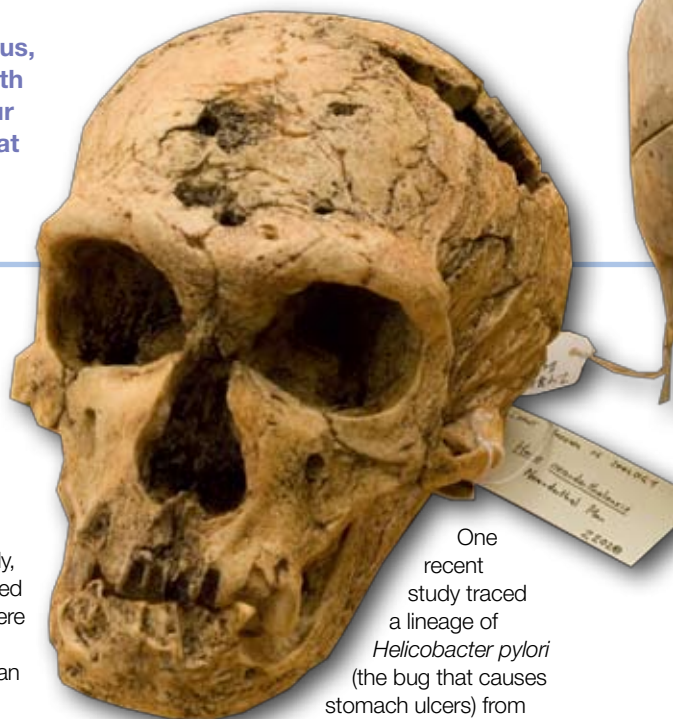
Following individual genes down more than a few human generations is next to impossible. Egg and sperm are each made in a specially orchestrated cell division, and their paired genes – one copy from each parent in the previous generation – are thoroughly shuffled before their single sets of chromosomes are assembled. These then combine to create a new person's genome. That happens in every round of reproduction, so clear lines of descent get very blurred.

Both sexes, though, contribute portions of DNA that escape this mixing-up. Males have a single copy of the genes on their sex-determining Y chromosome. And, as explained on page 8, everyone has lots of copies of the mitochondrial genes, which are only passed on from the egg. That means

there are two sets of genes that are normally passed cleanly from one generation to the next.

Analysis of both kinds has been used to track the history of human migration out of Africa. More recently, it has been extended to more detailed tracking of who may have been where and when.

More subtle traces of our past can also be found in the bacteria in our guts, and sampling shows they differ from place to place. Genome analysis of samples of a single species recovered from two different human populations can indicate when the populations became separated, assuming that the bacteria are undergoing a roughly constant rate of mutation.



One recent study traced a lineage of *Helicobacter pylori* (the bug that causes stomach ulcers) from populations scattered across the Polynesian islands of the Pacific Ocean. The researchers compared this with a similar mapping based on a close analysis of the likely evolution of the more than 400 different languages spoken on the islands – which also change over time. The conclusions

Getting physical

Small genetic changes can mean big physical differences.

We are fascinated by difference, but humans are really genetically rather similar. The latest figures suggest that we are all (identical twins excepted) 99.6 per cent similar.

Many human characteristics vary continuously – there are people at many different points on the scale. Height is a typical example, and recent analysis suggests at least 50 different genes (as well as environment) affect how tall a person grows.

A few characteristics of appearance vary more simply. Red hair, for example, goes with having two altered copies of a gene known as *MC1R*, which is involved in making the pigment melanin. There is some evidence for a number of the other characteristics often said to be associated with red hair, such as a lower tolerance for pain. The *MC1R*



gene is related to pain receptor genes in the brain. However, redheads' reputation for 'fiery' temperament probably has more to do with the associations of the colour than with anything in the DNA.

Less obvious personal differences, such as having softer earwax, have been tied to a specific gene. Although softer earwax is harmless, the same genetic variation has

been linked with increased risk of breast cancer, so it could turn out to be a medically useful sign.

Meanwhile, genomics is uncovering more details of the small differences in genes that can make a big difference to your appearance. For example, the International HapMap Project helped to pin down a single gene out of the 170 involved in hair growth that

makes a big contribution to hair thickness in eastern Asian populations. The gene *EDAR* comes in a thick-hair variant, which must have appeared after ancestral Asian and European populations divided. The gene appears to affect the activity of a transcription factor – one of the many ways one gene can alter the regulation of another. This factor is important in fixing the final thickness of hair fibres.

FAST FACT

The tallest man in history was 8' 11" (2.72 m). The current world's shortest man is 2' 5.37" (74.61 cm).





A modern skull (right) and a cast of a prehistoric skull (left). Extracting DNA that is of sufficient quantity and quality to be sequenced from ancient bones can be difficult, because of decay and DNA contamination. However, researchers announced in 2009 that they have sequenced the entire genome of a Neanderthal.

were a close match, and indicate strongly that the gradual colonising of the vast space of the Pacific spread out from what is now Taiwan, starting around 5000 years ago.

Putting the Y into you

What the Y chromosome tells us about our past.

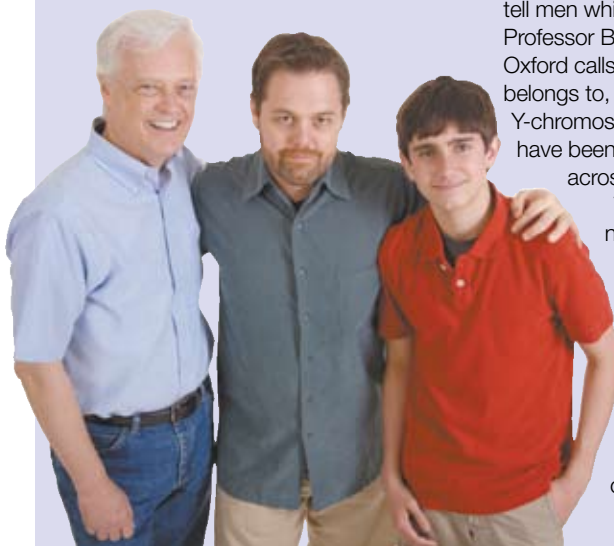
The male Y chromosome contains more DNA than the mitochondrion – around 50 million base pairs compared with 16 000. That difference provides greater scope for variation, allowing more detailed analysis of how DNA markers have been carried down the generations.

Such analysis can even shed light on aspects of individual ancestry, as well as tracking population movements. As many countries share our custom of passing surnames from father to son, DNA results

can be compared with old-fashioned genealogies and used to see whether two people with the same surname are genetically related. More remarkably, DNA profiles from a crime scene might be useful in predicting possible surnames of men who were there. A team led by Professor Mark Jobling at the University of Leicester is testing out how useful this could be, and whether it might raise unjustified suspicions.

Other analysis of the Y chromosome can tell men which one of several 'clans' – as Professor Bryan Sykes of the University of Oxford calls them – their Y chromosome belongs to, from the analysis of just 15 Y-chromosome markers. So far, 18 clans have been identified, varying in frequency across different parts of the world.

This sort of work is drawing new maps of countries, such as the UK, which have experienced repeated waves of migration. It can show, for example, how Viking place names and the surnames derived from them in parts of northern England go along with Viking-descended Y chromosomes.



Jeanel Norvell/Stockphoto

In your genes?

How much of 'me' is determined by my genome?

Our personality and how we behave are undoubtedly influenced by our genome. But there are real difficulties in researching the links. A lot of the possible traits you might study

– such as optimism, aggression or liking for novelty – can be hard to define. And, as with complex physical traits, genomic analysis typically finds many DNA variations that have weak associations with the behaviour in question.

Geneticists who work in the area seem to have a persistently

optimistic trait, though, continuing to work on the origins of a whole range of behaviours, including abilities such as reading and maths.

They continue to find new twists in the connections between genetics and life histories. For example, there is evidence from psychological studies that children's school performance is influenced by how smart they think they are – their 'self-perceived ability' (SPA). In 2009, Robert Plomin and colleagues at King's College London, who work with a large database of twins, reported that SPA is highly heritable. So while there is still a self-fulfilling prophecy involved, it may be at least partly due to some genetic influence on self-belief, rather than solely the messages children get from teachers or parents about their ability.

A different twist appears in

studies of the influence of a gene active in the brain that makes the protein monoamine oxidase A (MAOA). The gene affects levels of an important neurotransmitter, one of the chemicals that transmits signals across the gaps (synapses) between nerve cells. One rare form of the MAOA gene was implicated in the origins of antisocial behaviour in studies in the 1990s.

Later work found that children who had been ill-treated by their parents coped better if they had the normal, more active form of the gene than children with the variant form. Antisocial behaviour was linked to both ill-treatment and the rarer variant gene.

So perhaps the genetic influence is more to do with coping with some kinds of stress, rather than any more direct link with behaviour.



SEEING RED

Is red hair dying out? Do your bit to save this characteristic in the Wellcome Trust-funded game *Ginger Dawn*.

www.wellcome.ac.uk/bigpicture/genes

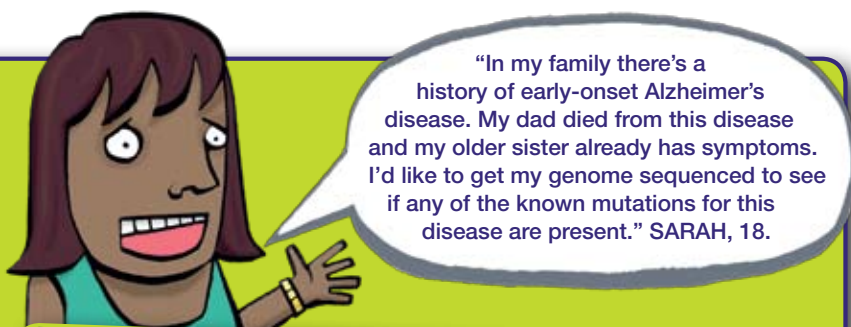
ON THE WEB

Your genome, your rights?

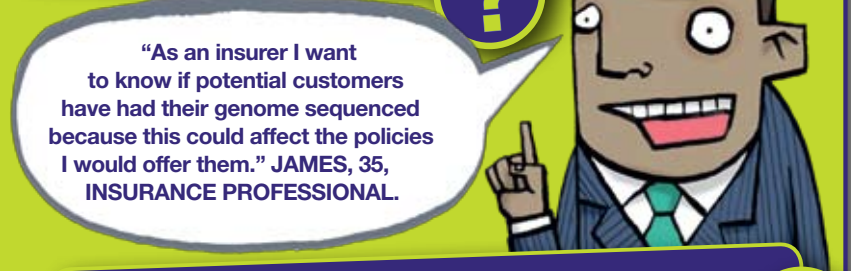
Genetic advances bring great potential for health, but they also raise questions around privacy and who should have access to personal information such as genetic data. Take a look at the made-up scenarios below to explore two issues that are far from black and white.

1. Insurance and genetic information

For some genetic conditions, there are predictive tests that can be used to show whether someone is likely to develop the disease. Potentially, whole-genome testing could give you a much wider view of your possible future health. There is some way to go until these tests are widely available, affordable and applicable to many common diseases, but imagine a future where this is the case, and think about the questions below. We've used the insurance industry as an example, but concerns about misuse of genetic data apply to many groups, including employers, governments and journalists.



Do you think Sarah should be allowed to get her genome sequenced?



Do you think it's right that insurance companies have access to potential customers' genome sequences before considering whether to offer them insurance?



Would you have your genome sequenced and allow scientists to have access to the information for the good of medical science?



CASE STUDY: HUNTINGTON'S DISEASE

In 2007, the Association of British Insurers reported that 132 people applying for insurance disclosed results for genetic tests for Huntington's. Of these, 108 had negative (i.e. 'normal') results and 19 had positive results, meaning that they carry the relevant mutation.

The table shows the outcomes for the 19 people whose tests showed that they carry the mutation for Huntington's.

Insurance decision	Number of people
Accepted at ordinary rates of insurance	2
Accepted with higher premiums or revised terms	14
Declined insurance	3

What regulation is there?

In May 2008, the Genetic Information Non-discrimination Act (GINA) was passed by the US Congress. Other countries have legislated against this kind of discrimination, including France, Finland and Sweden.

In the UK, there is a voluntary, rather than legislative, system in place. In 1999, the Genetics and Insurance Committee (GAIC) was established by the government. This committee oversees insurance companies to ensure that they comply with a code of conduct published by the Association of British Insurers. This code includes a moratorium on the use of genetic tests for insurance purposes. The moratorium lasts until 2014, and a review is planned for 2011.

The GAIC asks insurers to show the clinical, scientific and actuarial (relating to the statistics of risk) relevance of using the results of predictive genetic tests in working out insurance policy premiums. At the moment, the only test insurers are allowed to use is that for Huntington's disease, for life insurance policies over £500 000. Huntington's is a rare and incurable genetic disease that causes gradual deterioration of the brain. It is caused by carrying a faulty copy of the *Huntingtin* gene.

More broadly, in the UK, issues surrounding developments in human genetics are overseen by the Human Genetics Commission, a government advisory body. It includes a monitoring group for genetic discrimination.

With this in mind, what do you think will happen with such regulation in the future? Will more genetic tests be approved for use by insurers?



2. Genetic testing before birth

Prenatal genetic testing can be used to detect mutations in an embryo or fetus that are linked to particular conditions. Read the text below and study our imaginary headlines to see what you think about the questions that follow.

Prenatal genetic testing

Parents at high risk of having a child with a genetic condition will be given genetic counselling and will then be offered prenatal genetic tests. A limited number of diseases can be detected in this way, including sickle-cell disease, muscular dystrophy and cystic fibrosis. Prenatal genetic testing is also used to detect chromosome disorders such as Down's syndrome.

Tests for a number of other conditions are in development, including for autism. Some people argue that we should do all we can to rid society of any debilitating condition. Others think that tests for something such as autism would discriminate against 'differently abled' people, producing a society that's intolerant of difference.

Look at the headlines below to see how two papers tell the story in very different ways: the first more positive, the second less so. With these in mind, think about your own views and answer the question that follows.

Prenatal test for autism would deprive world of geniuses

Gene test hope for autism

Who do you think should be offered prenatal genetic testing? For what kinds of condition?



Preimplantation genetic diagnosis

This is a particular kind of prenatal testing. Embryos are created by *in vitro* fertilisation, then one cell is taken from each and tested for the presence of a particular genetic mutation. Embryos are then selected accordingly. This procedure can help couples of whom one or both carry disease-causing mutations to have unaffected children. Preimplantation genetic diagnosis can be used to test for over 50 conditions.

In 2008, preimplantation genetic diagnosis was used to help a woman with a family history of breast cancer to conceive a baby free of a mutation that causes early-onset, hereditary breast cancer.

Look at the headlines below to see how two papers tell the story in very different ways: the first more positive, the second less so. With these in mind, think about your own views and answer the question that follows.

Baby to be born free of breast cancer after embryo screening

Genetic embryo screening: where will it end?

Do you think that preimplantation genetic diagnosis should be used to screen embryos in this way?



Preimplantation tissue typing

This allows parents to select an embryo that could become a 'saviour sibling', a brother or sister who can donate 'matched' stem cells to a sick sibling affected by illnesses such as leukaemia. Here, parents select an embryo for implantation that has the same tissue type (i.e. that is matched for human leukocyte antigen) as the sick child.

The Human Fertilisation and Embryology Authority (HFEA) licenses preimplantation tissue typing (to produce 'saviour siblings') on a case-by-case basis in the UK. The first such licence was granted in 2002.

Look at the headlines below to see how two papers tell the story in very different ways: the first more positive, the second less so. With these in mind, think about your own views and answer the question that follows.

My little brother was born to save my life

Concern over 'spare part' babies

Do you think it's fair to select an embryo in this way to save a sibling?



What regulation is there?

Prenatal diagnosis has no special legislation apart from that covering termination of pregnancy (abortion). The decision about whether to have a test during pregnancy is made between the woman, her family and the doctors looking after her.

Genetic tests during pregnancy are available on the NHS if there is a family history of a condition, or if they are offered as part of a population screening programme. The UK National Screening Committee decides which conditions are screened for. While some tests may be available through private clinics, there is no test to screen all of a baby's genes.

The Human Fertilisation and Embryology Authority oversees all *in vitro* fertilisation procedures and research on embryos in the UK.

Sex selection is only allowed for medical reasons, for instance to avoid having a child affected by an X-linked disorder such as Duchenne muscular dystrophy. Such diseases, in which mutated genes are carried on the X chromosome, more often affect males (who have one X chromosome) than females (who have two).

Question to discuss: Who should decide what prenatal genetic tests and screening are carried out?



LAW OUT OF ORDER?

In 1995, the Police National DNA Database was launched in England and Wales to allow the police to store DNA profiles. Do you think it should be compulsory for everyone's DNA to be stored on the database? Take a look at our exclusive online lesson plan and video to explore the issues around this.



ON THE WEB

www.wellcome.ac.uk/bigpicture/genes

Real voices

Progress in genetic research affects all of our lives, but what's it like to face these issues every day in your work or personal life? We talk to a genetic counsellor, a research scientist and a man who carries the mutation that causes Huntington's disease to find out.



Anaar Sajoo

What do you do?

I am a Principal Genetic Counsellor, and have been working in the field for 15 years.

Why did you become a genetic counsellor?

I did a science degree at university. I liked the science part but I didn't want to spend all my time in the lab, and wanted some human contact, so I trained as a genetic counsellor.

What does your job entail?

Some genetic counsellors specialise, but I cover most areas, including prenatal diagnosis, cancer genetics and neurological conditions. The diseases I counsel on include cystic fibrosis, muscular dystrophy and Huntington's disease.

About half of what genetic counsellors do is to study families with lots of cancer. We test for genetic mutations that could be increasing a family's risk of linked cancers – we find these in around 20 per cent of the families. This can be heartening work as this kind of knowledge can help subsequent cancers be detected early, and there are often therapies available.

How do you break bad news to patients?

People often think that giving bad news is the hardest part of my job, but it's not something I do that often. Also, what we think of as 'bad' may not be so for the patient. When we do have 'life-and-death' news to give we plan with the patient how they'd like to hear it – we can phone them before we see them in person, for example.

How has genetic testing changed over your career?

We don't have cures for genetic conditions yet, but there's been a lot of progress – such as increased life expectancies, better treatments and new ways to test – so we have to stay up to date. We are guided by the family, who deal with the condition first-hand, every day. Genetic counsellors translate complex medical and genetic information. Our advantage lies in listening to what they are going through, and offering a combination of support and genetic information.

For information on a career in genetic counselling, see: www.agnc.org.uk/howtobecomeaGC.htm



Elizabeth Murchison

What do you do?

I'm a postdoctoral researcher at the Wellcome Trust Sanger Institute, Cambridge, working at the Cancer Genome Project.

Why did you become a scientist?

I was really interested in biology as I was growing up. I became aware of the Human Genome Project as I was finishing school – I was quite disappointed as it was just ending and I thought there'd be nothing left to do. Then I discovered the world of molecular biology and genetics and realised that the Human Genome Project was just the beginning. After completing a biology degree in Australia I got my PhD in the USA.

What does your job entail?

Cancer is caused by genetic changes – mutations – in a person's DNA. For the Cancer Genome Project, we're looking at genomes taken from cancerous cells to try and work out how these genetic changes could contribute to cancer development and cancer evolution.

We take tumour tissue and normal tissue from the host and then sequence the genomes of both. We can't use a standard human genome to compare the cancer genome to, as every individual is much more different to each other than an individual is to their cancer.

What cancers do you study?

At the Cancer Genome Project, some researchers focus on specific kinds of cancer, others screen 'their' gene through lots of different cancers. Most research is on human cancers, but I'm the exception: I study transmissible cancers in Tasmanian devils and dogs. Transmissible cancers are spread by the physical transfer of cancer cells. They are very unusual and, as I grew up in Tasmania, I wanted to do something to help the Tasmanian devil.

How has sequencing changed over your career?

The progress of sequencing technology is just incredible – it's moving so quickly that we're always fighting to stay on top of it. My job is a lot of work and it can be hard to take holidays, but I'm so excited by the new discoveries I come across in the data we generate. It's a huge thrill to be the first person to see something biological, and there's real excitement in the air where I work.

For more information on the Cancer Genome Project, see: www.sanger.ac.uk/genetics/CGP/

WHAT IT'S LIKE FOR ME

Meet other young people living with genetic conditions and find out more about the science involved, with our exclusive online interviews.

www.wellcome.ac.uk/bigpicture/genes

ON THE WEB



Mediastock/Photolibrary

Get online

Everything you need to take your lessons to the next level at the click of a mouse.



ON THE WEB

Although you've nearly reached the end of the magazine, the *Big Picture* action doesn't stop there. Go to www.wellcome.ac.uk/bigpicture/genes to browse exclusive online articles, videos and other content, which take you further into the topics covered in this issue.

Find out what researchers can learn by sequencing the genomes of pathogens, watch our video about how the discovery of a genetic mutation has led to a treatment that's transformed the life of children with neonatal diabetes, and do your bit to save ginger-haired people in the game *Ginger Dawn*. Also online is our exclusive curriculum-matched lesson plan and video to help you and your students discuss the issues around the Police National DNA Database.



Matt Ellison

What do you do?

I'm 21 and a full-time carer for my dad, who has Huntington's disease.

What is Huntington's disease?

Huntington's is a genetic disease that causes gradual deterioration of the brain. It can affect all ages; there's a juvenile form that occurs in children as young as five.

My dad first got symptoms when I was at primary school. They were very mild at first, and it seemed like a good thing to me because he was home all the time and I saw a lot of him. When I was 12 he started to get really bad. He's 54 now and in the late stages: he's unable to talk, walk or eat.

How has Huntington's disease affected your life?

It is very difficult growing up in a family with Huntington's disease. My dad's illness means that I found out about Huntington's early in life, and my mum was very open about everything. I didn't really think about it in terms of my own risk until I was 18 – I don't know why, it just hit me then that I needed to face it.

How does the testing process work?

I asked my GP to refer me to a genetics team. They see you every month for six months to check how you're coping with the process, then you get tested. In June 2008 I found out I was positive for the mutation. It was a surreal day, really strange. I was glad to find out, but at the same time it was terrible news that I didn't want to hear.

How has your diagnosis changed you?

After I was tested I started fundraising for the Huntington's Disease Association. Running was completely new to me, but since June 2008 I've run four marathons and have two more to do.

Doctors can't predict when the disease will start to affect me, but it usually occurs earlier in the next generation if it's passed down the male line – so I've probably got 10–15 good years left. With Huntington's you either sit back and let it defeat you, or you stand up and fight back a bit. It took me a few months to get back on my feet, but once I did I realised that I can take a lot of strength from it.

For Matt's marathon progress, see: www.justgiving.com/mysorelegs

Sadly, Matt's dad died three weeks after this interview. Matt and his family wanted this article to be published.

Opinions wanted!

Your chance to be on the *Big Picture* Teachers' Advisory Board.

Would you like to be involved in the development of *Big Picture* and have your say about future topics and content? If you've got ideas about what resources you'd like us to produce, or thoughts about how we could support science teachers better, then get in touch and let us know.

The *Big Picture* team is setting up a **Teachers' Advisory Board** and we're looking for science teachers from across the UK to join us. The time commitment will be minimal, and we'll provide incentives.

For more information or to register your interest, please email bigpicture@wellcome.ac.uk, subject line 'Teachers' advisory board'.



MBPHOTO/Stockphoto

Education editor: Stephanie Sinclair

Editor: Chrissie Giles

Writers: Jon Turney, Chrissie Giles

Illustrator: Glen McBeth

Graphic designer: Paulo Estriga

Project manager: Jennifer Trent Staves

Advisory board: Jess Buxton, Dave Carr, Nan Davies, Rachel Iredale, Jane Itzhaki, Mark Jobling, Rhona Knight, Julian Parkhill, Michael Reiss, Bronwyn Terrill

All images, unless otherwise indicated, are from Wellcome Images.

The Wellcome Trust is a charity whose mission is to foster and promote research with the aim of improving human and animal health (a charity registered in England, no. 210183). Reflecting the profound impact today's research will have on society, the Wellcome Trust also seeks to raise awareness of the medical, ethical and social implications of research and promote dialogue between scientists, the public and policy makers.

ISSN 1745-7777

© The trustee of the Wellcome Trust 2010.

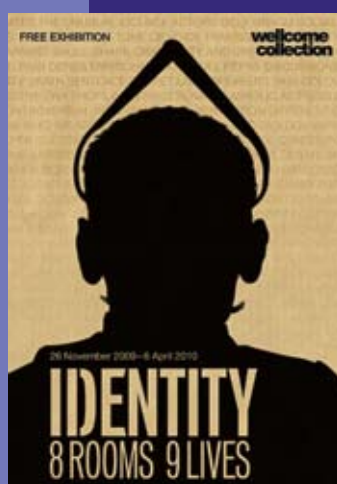
This is an open access publication and, with the exception of images and illustrations, the content may unless otherwise stated be reproduced free of charge in any format or medium, subject to the following conditions: content must be reproduced accurately; content must not be used in a misleading context; the Wellcome Trust must be attributed as the original author and the title of the document specified in the attribution.

PU-4638/14K/12-2009/PE

- What is genome sequencing, and how and why is it done?
- How do our genes affect our personality, appearance and health?
- What can studying DNA tell us about human history?
- How similar are we to other organisms?
- How should genetic information be used?

In 2000, the working draft of the human genome was unveiled. Published a decade later, *Big Picture: Genes, genomes and health* explores the developments in sequencing and genetic research made since this landmark event. Look inside to discover how technological developments are raising new questions around the regulation of and access to genetic information.

Explore what we know now about the human genome and how it compares with those from other living things. Find out how the genetic information we carry in nearly all of our cells can not only give an insight into our past, but is also beginning to reveal clues about our lives and health in the future.



The Identity Project is a nine-month season of activity from the Wellcome Trust, starting in November 2009, including a major exhibition and diverse events presented in Wellcome Collection, plus exhibitions, live events and films at other venues across the UK. The season explores scientific and social perspectives of identity – historical and contemporary – to encourage debate and discussion and to ask how well we will ever be able to know ourselves.

www.theidentityproject.org.uk

BIG PICTURE SUBSCRIPTIONS

SIGN UP TO RECEIVE REGULAR FREE COPIES

Sign up to receive a free single copy or multiple copies of each issue. Or order a specific issue, or copies of any of the previous issues in the series – all free of charge.

To sign up, visit
www.wellcome.ac.uk/bigpicture/order

Or contact us by telephone, email or post at:

T +44 (0)20 7611 8651
E publishing@wellcome.ac.uk

Big Picture series
Wellcome Trust
FREEPOST RLYJ-UJHU-EKHJ
Slough SL3 0BP



FEEDBACK

If you have any comments on the *Big Picture* series, please email bigpicture@wellcome.ac.uk.



This document was printed on material made from 25 per cent post-consumer waste & 25 per cent pre-consumer waste.

